A simple two-component model of the electrically evoked compound action potential in the human cochlea

Lai, W K ; Dillier, N

Abstract: Neural response telemetry (NRT) permits in situ intracochlear recordings of the electrically evoked compound action potential from the auditory nerves using scala tympani electrodes. The recorded NRT waveforms can generally be categorized under either single positive peak or double positive peak waveforms. This is similar to the observations from Stypulkowski and van den Honert, who suggested that the double peak complex arises from two components that could be axonal and dendritic in origin, respectively. Using a simple mathematical model which linearly combines two separate waveforms similar in shape but differing in amplitude and latencies, it was possible to simulate the various NRT waveform categories. The simulation results support the view that the two waveform components originate from dendritic or axonal processes and implies that the shape of the response waveform may provide information about the degree of neural survival in the stimulated cochlea. This information could be useful for determining optimal speech coding parameters for cochlear implant users on an individual basis.

DOI: https://doi.org/10.1159/000013899

Original Paper

Audiology & Neuro-Otology

Audiol Neurootol 2000;5:333–345

Received: June 21, 1999
Accepted: April 4, 2000

A Simple Two-Component Model of the Electrically Evoked Compound Action Potential in the Human Cochlea

Wai Kong Lai Norbert Dillier

ENT Department, University Hospital, Zurich, Switzerland

Key Words

Evoked potentials · Cochlear implant · Neural response telemetry · Mathematical modeling

Abstract

Neural response telemetry (NRT) permits in situ intracochlear recordings of the electrically evoked compound action potential from the auditory nerves using scala tympani electrodes. The recorded NRT waveforms can generally be categorized under either single positive peak or double positive peak waveforms. This is similar to the observations from Stypulkowski and van den Honert, who suggested that the double peak complex arises from two components that could be axonal and dendritic in origin, respectively. Using a simple mathematical model which linearly combines two separate waveforms similar in shape but differing in amplitude and latencies, it was possible to simulate the various NRT waveform categories. The simulation results support the view that the two waveform components originate from dendritic or axonal processes and implies that the shape of the response waveform may provide information about the degree of neural survival in the stimulated cochlea. This information could be useful for determining optimal speech coding parameters for cochlear implant users on an individual basis.

Introduction

With the advent of increasingly sophisticated multichannel cochlear implant systems able to implement more than one stimulation paradigm or coding strategy, the task of providing the user with individually optimized programs in their speech processors has become more complex. For instance, the various stimulation paradigms presently in common usage – e.g. MPEAK (Multi Peak) [Skinner et al., 1997; Cafarelli Dees et al., 1995], SPEAK (Spectral Peak) [Whitford et al., 1995; Skinner et al., 1994], ACE (Advanced Combination Encoder) [Vandali et al., in press], CIS (Continuous Interleaved Stimulation) [Boex et al., 1996; Kiefer et al., 1996; Wilson et al., 1993] – each produce quite different sound percepts, often requiring a certain period of familiarization before they can be properly evaluated. Consequently, the process of selecting an optimal stimulation paradigm can be rather time-consuming. The fitting procedures will also become more complex as an increasing number of parameters (e.g. the rate of stimulation, the number of active channels) need to be optimized in order to achieve good individually fitted speech processor programs. The problems are further compounded for the ever-increasing number of very young children being implanted who are less likely to be able to provide accurate subjective feedback about the settings of the fitting parameters. Objec-

KARGER

Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

1420–3030/00/0056–0333$17.50/0

Accessible online at:
www.karger.com/journals/aud

Wai Kong Lai, PhD
ENT Department, University Hospital
Frauenklinikstrasse 24
CH–8091 Zürich (Switzerland)
Tel. +41 1 255 58 01, Fax +41 1 255 44 24, E-Mail lai@orl.usz.ch

tive measures are therefore desirable to assist with the optimization process by narrowing down the initial choices of parameter settings for the individual user.

One prospective approach would be to examine the behavior of the eighth nerve to electrical stimulation in order to deduce the capacity of the implant user's auditory pathway to transmit the stimulus information. Methods ranging from direct measurements involving single auditory nerve fibers to measurements of the compound action potentials as well as early and late auditory responses using various electrophysiological techniques have been employed to achieve a better understanding of how the auditory periphery functions. Methods examining auditory nerve potentials have the advantage that they help to characterize the peripheral behavior more accurately. However, access to the peripheral nerves necessitates special equipment. Very often, the subject also needs either to remain very quiet or to be sedated during the measurements which would then still require averaging over hundreds or thousands of sweeps to extract the desired signal from the accompanying measurement noise.

Recently, a new and convenient method for measuring the intracochlear evoked compound action potential (ECAP) has become available with the Nucleus 24 multichannel cochlear implant system [Carter et al., 1995]. Called neural response telemetry (NRT), this measurement method was developed based on the intracochlear ECAP measurement techniques used by Brown et al. [1990]. The Nucleus 24 implant features a number of back-telemetry options, one of which is able to measure the evoked whole-nerve action potentials using the implant's own intracochlear electrodes as recording electrodes. Integrated within the implant are voltage-sampling, encoding and back-transmission circuitry. Specially written NRT software supports communication with the implant to perform entire ECAP measurements using the standard clinical equipment without additional hardware.

The proximity of the intracochlear recording electrodes to the neural interface means that the susceptibility of the signal to noise is considerably reduced (and so too the amount of averaging or recording time required to obtain a clear recording of the neural response). This improved noise immunity also means that the subject does not have to be sedated for a measurement session. The measurements can indeed be conducted with an awake subject who may simultaneously be occupied with sedentary activities such as knitting, reading or watching a (silent) video.

The measured neural response waveform provides a record of the behavior of the auditory nerve fibers in response to electrical stimulation. By examining various aspects of the response waveform systematically, it may be possible to deduce useful information about these auditory nerve fibers. For instance, the size of the response amplitude could be related to the population density of the responding neurons, or their temporal response properties could be related to the capacity of those nerve fibers to transmit information encoded in the rate of stimulation. Another aspect of the response waveforms, and one which this paper will concentrate on, is the shape of the waveform itself, which could provide useful insights into the way these neural responses are generated.

Stypulkowski and van den Honert [1984] reported that the ECAP of the intact auditory nerve from cats exhibits two peaks. Following cochlear laminectomy where the basilar membrane (and presumedly the peripheral processes) was removed, the resulting action potentials showed only the earlier of the two peaks. They then postulated that the two peaks represented action potentials initiated from either the axon of the cells or from the dendritic processes. These results were based on the extreme cases where there was either a fully intact cochlea or the entire periphery was missing, resulting in a total absence of the later positive peak. In pathological cochleae, the proportion of surviving axonal and dendritic processes will vary from patient to patient, as well as from site to site along the cochlea. Different proportions of surviving axonal and dendritic processes could foreseeably give rise to a range of response waveforms. In order to establish whether it is possible for such a mechanism involving two sites of origin of neural activity to give rise to the range of clinically observed NRT waveforms, a mathematical model was constructed.

This paper consists of two parts. In one part, the NRT measurement system and the actual measurement data are briefly described. A categorical scale for classifying the response waveforms is then presented. In the other part, a simple first-order mathematical model which generates compound waveforms similar to the NRT response waveforms using two-component waveforms is presented. Based on the simulation results using this model to reconstruct the observed NRT waveforms, a simple method is proposed by which NRT waveforms may be interpreted to identify the relative contributions from these two components. Finally, the implications of these results for optimizing the fitting procedures of the cochlear implant users' speech processors will be discussed.

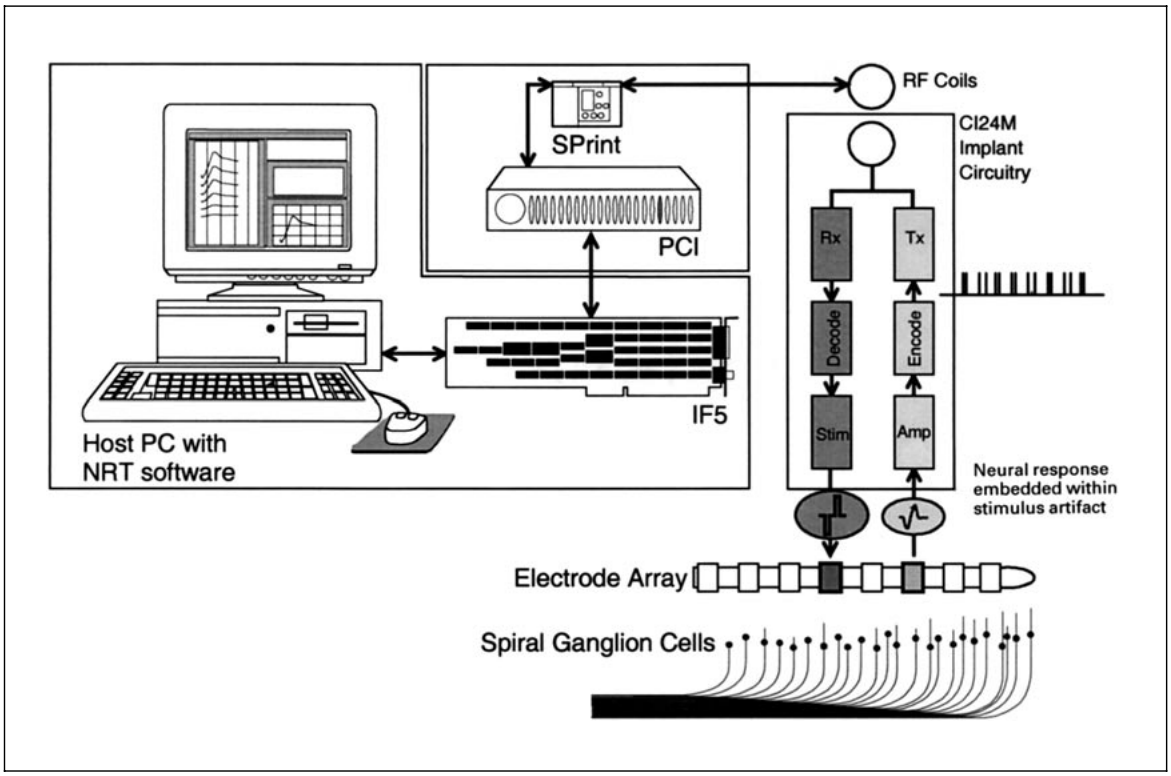

Fig. 1. The NRT measurement system consisting of a host PC with custom NRT control software which communicates with the CI24M implant via the standard Nucleus clinical equipment (IF5 interface card, PCI processor control interface and SPrint speech processor). Encoded stimulus and recording parameters are transmitted via radio frequency (RF) link and are received (Rx) by the implant. After the desired stimuli have been presented, the resultant action potential from the activated nerve fibers is recorded by the measurement amplifier, encoded and then transmitted (Tx) back to the SPrint processor over the radio frequency link. The recorded data are processed by the NRT software and stored in the host PC.

Method

The NRT Measurement System

Principles of Operation

The NRT measurement system illustrated in figure 1 comprises the standard Nucleus 24 clinical system (i.e. IF5 interface card, PCI processor control interface and SPrint speech processor) together with custom NRT software [operating under the Windows (3.1, 95 and 98) environment] which controls the stimulation and measurement parameters. These parameters are encoded by the SPrint speech processor as a sequence of radio frequency bursts for transmission across the skin to the implant's receiver/stimulator. This encoded information received by the implant is then decoded and used to (1) configure the implant's ECAP measurement circuitry and (2) stimulate the selected electrodes accordingly. The resultant ECAP (as well as any accompanying stimulus artifact) is then sampled at the specified measurement electrode and encoded for transmission back across the skin to the SPrint speech processor via the radio frequency coils. The NRT software communicates with the SPrint speech processor to capture, process, store, display and analyze these measurement data on the host personal computer (PC).

Subtraction Paradigm

To separate out the relatively small neural action potentials from accompanying stimulus artifacts in the ECAP recording, the NRT software implements a forward masking artifact subtraction paradigm. Also referred to as the subtraction paradigm, it is a modified version of the method described by Brown et al. [1990] for cochlear implant users which was in turn adapted from the method originally described by Charlet de Sauvage et al. [1983].

The subtraction paradigm (illustrated in fig. 2) essentially takes advantage of adaptation effects to separate the neural response from the stimulus artifact. Using three different stimulus intervals A (presenting the probe alone), B (presenting a masker followed by the probe) and C (presenting the masker alone), the neural response can be separated from the stimulus artifacts arising from both masker and probe by subtracting the recordings from the three respective intervals according to A – (B – C).

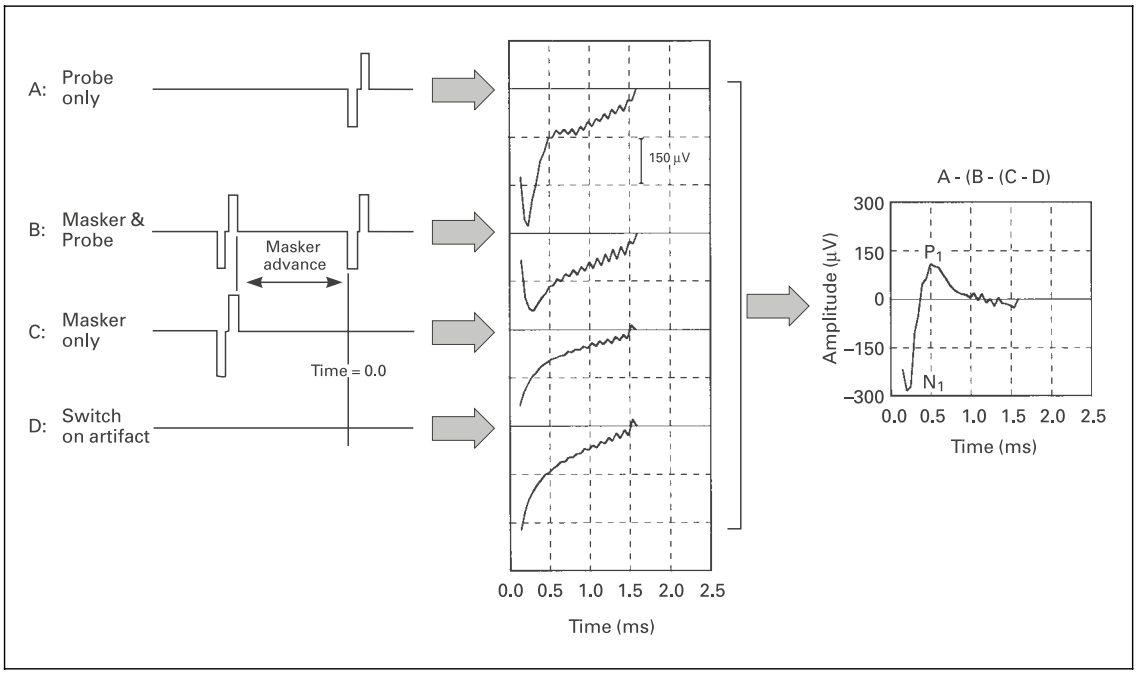

Fig. 2. The subtraction paradigm as implemented by the NRT system to separate out the relatively small ECAP or neural response from the accompanying stimulus artifacts. Note that the ordinate is plotted positive upward for the individual recordings from the four stimulus intervals.

A further artifact to be accounted for is caused by the measurement amplifier itself and is produced every time the amplifier is switched on to perform a measurement. This switch-on artifact can be measured using a fourth interval D which presents both masker and probe at minimal amplitude, simulating a stimulus interval in order to initiate a measurement, but producing no accompanying stimulus artifacts. This artifact is eliminated from the recordings by subtraction according to [A – (B – C)] – D to yield the desired neural response.

Figure 2 also shows a typical neural response waveform with a negative peak N_1 followed by a positive peak P_1 before the response waveform settles back to zero. The response amplitude is measured simply as the amplitude difference between the P_1 and N_1 peaks.

Stimulation and Recording

An ECAP measurement is initiated by first presenting the biphasic stimulus on a selected electrode pair. The stimulus is followed by a blanking delay of programmable duration before the amplifier is enabled to perform the voltage measurements. The blanking delay is needed to prevent the amplifier from being driven into saturation by the stimulus artifact and is also critical in determining the time window available to capture the response waveform. A sufficiently long delay is necessary to avoid nonlinear behavior (due to saturation or otherwise) of the amplifier caused by the stimulus artifact. Too long a delay, however, would result in a time window which misses capturing important early features of the neural response such as the first negative peak. Thus, the measurement parameters, particularly the delay, need to be carefully adjusted to optimize this measurement time window. Further details on the optimization of the measurement parameters have been documented by Abbas et al. [1999] and Lai [1999].

At the end of this delay, the blanking is removed and the evoked action potential is then measured from another selected pair of electrodes as a series of voltage samples (32 samples at 20 kHz sampling rate) spanning a time window of 1.55 ms. It is these 32 voltage samples that are encoded and transmitted back first to the speech processor and then to the controlling NRT software on the host PC as the neural response.

Both stimulation and recording can be performed in either monopolar or bipolar configuration. Typically, stimulation in monopolar configuration (MP1 extracochlear reference ball electrode) is used, together with recording in monopolar configuration (MP2 extracochlear reference plate electrode). Bipolar stimulation and monopolar recording configurations can also be used. Bipolar recording configurations, however, generally yield poor recording results [Carter et al., 1995].

NRT Data Collection

NRT recordings were collected on a routine basis for all 60 Nucleus 24 implant users implanted at the ENT Department of the University Hospital in Zurich. No subjects were excluded on the basis of etiology or speech performance. Postoperative recordings were made for all 60 subjects. Intraoperative recordings were also made for 24 of the subjects.

Intraoperatively, NRT recordings were made for all 22 intracochlear stimulation sites. Postoperatively, a minimum of 5 stimulation sites (electrodes 20, 15, 10, 05 and 03) spread out along the intracochlear array were tested. The corresponding recording site was typically chosen to be two electrodes apical to the stimulation site (i.e. the neural response to stimulation on electrode E would be recorded from electrode E + 2). Where this was not possible (i.e. for stimulation presented on the most apical electrodes 21 and 22), the recording site was chosen to be two electrodes basal to the stimulation site. Stimulation or measurement sites indicated by the Nucleus WinDPS programming software to be either open (>20 kΩ) or short (<200 Ω) circuit were not used and were substituted by a neighboring site whenever possible.

Biphasic stimuli of 25 μs per phase were used for both masker and probe stimuli. The subtraction paradigm requires that there be sufficient adaptation due to the masker for effective artifact subtraction. A masker amplitude larger than the probe amplitude by 5 current level units was employed, with the masker preceding the probe by 500 μs. The loudest acceptable presentation level as judged by the subject was first determined for each stimulation site prior to making any recordings. Stimuli were never presented above the loudest acceptable presentation level, which, depending on the subject, may be either substantially higher than or close to the behavioral comfortable loudness level. The recording parameters can be adjusted to optimize the quality of the recorded waveform. If no neural response could be observed despite adjustment of the recording parameters, measurements would proceed to the next predetermined stimulation/recording site. Recordings averaged over 100 sweeps at either 35 or 80 Hz were made with the measurement amplifier gain set at 60 dB. When the lower gain of 40 dB was selected, the number of sweeps was increased to 200.

Altogether, NRT recordings were performed for a total of 785 stimulation/recording sites.

Mathematical Modeling of the Response Waveforms

A simple mathematical model based on the assumption that the measured neural response consists of only two components was constructed. The two components supposedly represent the neural response from the respective populations of axonal and dendritic processes that are activated, as postulated by Stypulkowski and van den Honert [1984]. This model uses a linear combination of the two components to produce the compound waveform. In the simulations, the relative amplitudes and latencies of these two components were then independently adjusted to see if this would give rise to similar waveform morphology as obtained using NRT.

To perform the simulations, the impulse response $r(t)$ of a second-order low-pass Butterworth digital filter (cutoff frequency 100 Hz, Nyquist frequency 1000 Hz) was sign-inverted for use as a first-order approximation of the neural responses from the dendritic periphery and from the axons, respectively. An exponential weighting function t^p, where p is the independent variable, was employed to alter the relative amplitudes of the positive and negative peaks (i.e. skewness) of this impulse response $r(t)$, while the overall amplitude of the waveform was set by a separate attenuation factor m. Thus, each response component was modeled by the function

$$m \times r(t) \times t^p \qquad (t \geq 0), \tag{1}$$

where m and p are both real numbers. Note that the sign inversion of $r(t)$ is reflected in a negative value of m.

According to the physiological model from Stypulkowski and van den Honert [1984], the axonal component should occur earlier than the dendritic component for a recording electrode placed on the auditory nerve. The two waveforms should differ in amplitude, skewness as well as the onset time. If the two (axonal and dendritic) components above were represented by the time functions $r_a(t)$ and $r_d(t + \delta)$, respectively, the combined response $R(t)$ would be given by

$$R(t) = m_a \times r_a(t) \times t^{n_a} + m_d \times r_d(t + \delta) \times (t + \delta)^{n_d} \quad (t \geq 0), \tag{2}$$

where m_a and p_a are the amplitude factors of the first (axonal) component, and m_d and p_d are the amplitude factors of the second (dendritic) component. The two functions $r_a(t)$ and $r_d(t + \delta)$ are similar except that the latter is delayed by time δ. The simulations were made by adjusting m_a, m_d, p_a, p_d and δ. Note that as the simulations involved digital filters, δ was implemented as an integer variable.

The simulations were conducted using MATLAB (from The MathWorks Inc., Natick, Mass., USA) to reconstruct the neural response waveforms.

Results

NRT Waveforms

From the total of 785 stimulation/recording sites tested, clear neural responses could be obtained for 665 sites (84.7%). Of these, the NRT measurement data exhibited a variety of response amplitudes and response waveforms. The response waveforms were characteristic of evoked action potentials and may be broadly classified according to their morphology into two main categories, namely single positive peak (category I) responses and double positive peak (category II) responses. The single peak (category I) responses may in turn be further classified into three subcategories Ia, Ib and Ic. Examples of each category and subcategory are shown in figure 3 and are described in greater detail as follows.

Category Ia

This first subcategory of single positive peak responses represents the most commonly encountered and recognizable shape of the NRT response waveforms. It is characterized by a distinct negative peak N_1 followed clearly by a positive peak P_1 of smaller amplitude. The N_1 peak typically occurs around 300–400 μs and the P_1 peak at around 600–700 μs. In a limited number of cases, the P_1 peak occurred as early as 400–500 μs.

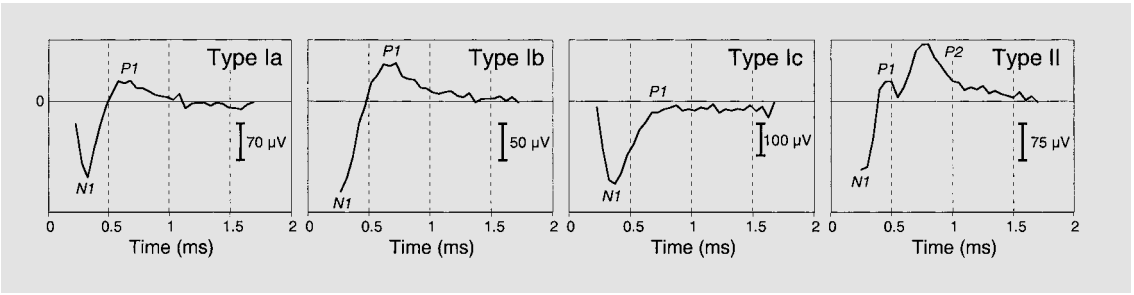

Fig. 3. Examples of the four NRT waveform categories. The stimuli consisted of biphasic pulses of 25 µs per phase. Monopolar stimulation and recording configurations were used, with the recording site always two electrodes away from the stimulation site.

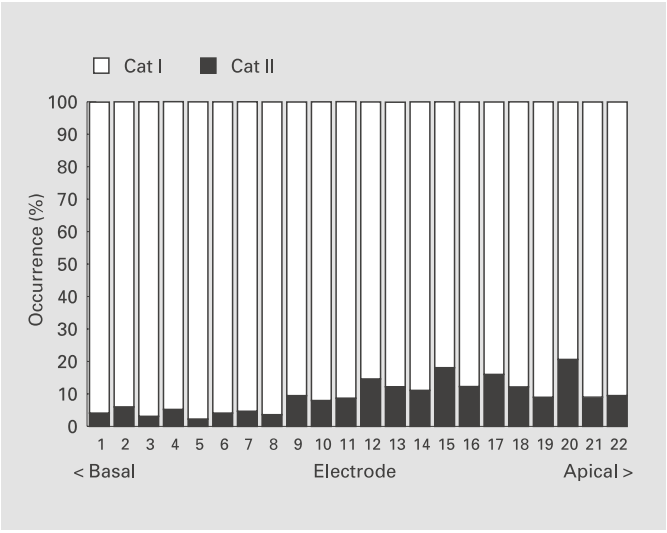

Fig. 4. Proportion of occurrences of category I and category II responses across all 22 stimulation/recording sites. Electrode 1 is the most basal site and electrode 22 the most apical one in the array.

Category Ib

This second subcategory of responses resembles the former, except that the N_1 peak is not visible; its presence is implied only by the rising trajectory of the response waveform. The N_1 peak appears to have occurred earlier than can be captured within the recording time window. Together, category Ia and Ib responses accounted for 82.2% of the waveforms obtained.

Category Ic

The third subcategory of single positive peak response waveforms resembles both category Ia and category Ib responses. This time, however, there is no clear P_1 peak visible although the trajectory from the negative peak is present (with or without a visible N_1 peak). The waveform appears to have a rising trajectory that asymptotes towards zero, giving the impression of a response with only a negative N_1 peak. These responses were less common (8.3%) than the former two categories of responses.

Category II

This category of responses differs clearly from the previous ones in that the response waveforms display two distinct positive peaks P_1 and P_2, both occurring after the negative N_1 peak. The N_1 peak is often not visible in these responses. Typically, the first peak P_1 occurs around 400–500 µs, and the second peak P_2 around 600–700 µs. In general, category II responses were relatively rare (9.5%).

Figure 4 shows the proportion of occurrences of the various response categories obtained across all 22 stimulation/recording sites. With the array divided equally into basal and apical segments (11 electrodes each), a χ^2 test showed no significant difference in the distribution of the category I responses between the two segments, implying that these responses were uniformly distributed across the array. Category II responses, on the other hand, were observed significantly (<0.1% confidence level) more often at the apical half of the array.

Modeling Results

Please note that for the purposes of these simulations, category Ib responses were regarded as essentially category Ia responses whose P_1 peak could not be captured with-

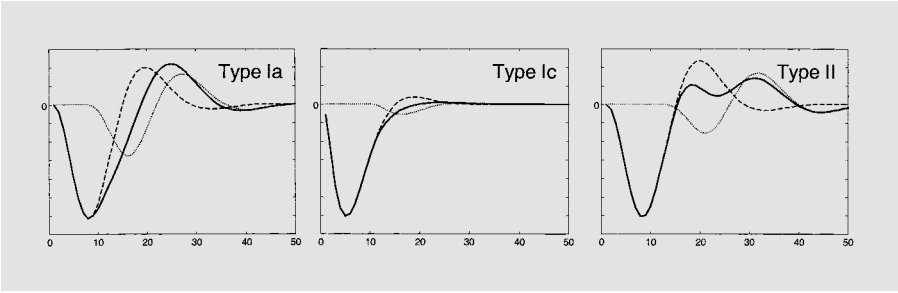

Fig. 5. Simulation results from linearly combining two component waveforms of different relative amplitudes and relative onset times. The dashed and dotted lines represent the $r_a(t)$ and $r_d(t)$ components, respectively, and the solid line the combined response waveform $R(t)$. Both the time (x-axis) and amplitude (y-axis) scales are in arbitrary units.

in the recording time window. Thus, the two categories were treated as one.

Using this simple linear model above, it was possible to produce compound waveforms that resemble category Ia, Ic and II response waveforms by adjusting the relative onset time and relative amplitudes of the two component waveforms, as demonstrated in figure 5. The values of the m_a, m_d, p_a, p_d and δ parameters in equation 2 used to reproduce these response waveforms are summarized in table 1. It should be noted here that the time axes of the diagrams shown in figure 5 are in arbitrary time units dependent on the sampling frequency used for the filter design. By selecting an appropriate sampling frequency, it is possible to match the temporal characteristics of the simulated waveforms to those of the measured NRT waveforms.

An interesting outcome of this simulation is that certain amplitude and delay combinations of the two components could actually produce a compound response waveform with only a single positive peak, similar to category Ia (and category Ib) responses. According to the mathematical model, these single positive peak responses consisted approximately of a 2:1 proportion of the $r_a(t)$ and $r_d(t)$ components with the $r_d(t)$ component appearing at the same time as the negative peak of $r_a(t)$. However, it must also be pointed out that single peak responses may also arise from the trivial case when only a single $r_a(t)$ or $r_d(t)$ component is present.

Category Ic responses which have no distinct positive peak appeared to consist mainly of an $r_a(t)$ component combined with a small $r_d(t)$ component of intermediate latencies.

The simulations of the double peak category II responses indicated a slightly greater proportion (approx. 3:1) of the $r_a(t)$ and $r_d(t)$ components compared to category Ia responses, as well as a larger latency difference between the two components. The simulation results suggest that it was the increased latency between the $r_a(t)$ and $r_d(t)$ components which gave rise to the double peak structure.

Table 1. Summary of model parameter values used to produce the simulation results for category Ia, Ic and II response waveforms as shown in figure 5

	m_a	p_a	m_d	p_d	δ
Category Ia	−1.54	1.8	−0.15	2.5	7
Category Ic	−28.6	0.25	−1.0	0.8	10
Category II	−1.0	2.0	−0.14	3.3	11

Discussion

Rationalizing the Single Peak Response Subcategories

Of the three subcategories of single positive peak category I responses, category Ib responses can easily be visualized as category Ia responses whose N_1 peaks have occurred earlier than can be captured within the measurement time window. For instance, when a large stimulus artifact is present, it is necessary to increase the blanking delay to avoid nonlinear behavior of the measurement amplifier when it is enabled. The increased delay could result in a measurement window that begins later than the actual occurrence of the N_1 peak. Another possible reason for not capturing the N_1 peak is when it has a very short latency (<110 μs), occurring before the minimum blanking delay of the measurement amplifier has elapsed and the amplifier can be enabled. Such an N_1 peak of very short latency may be interpreted as corresponding to very

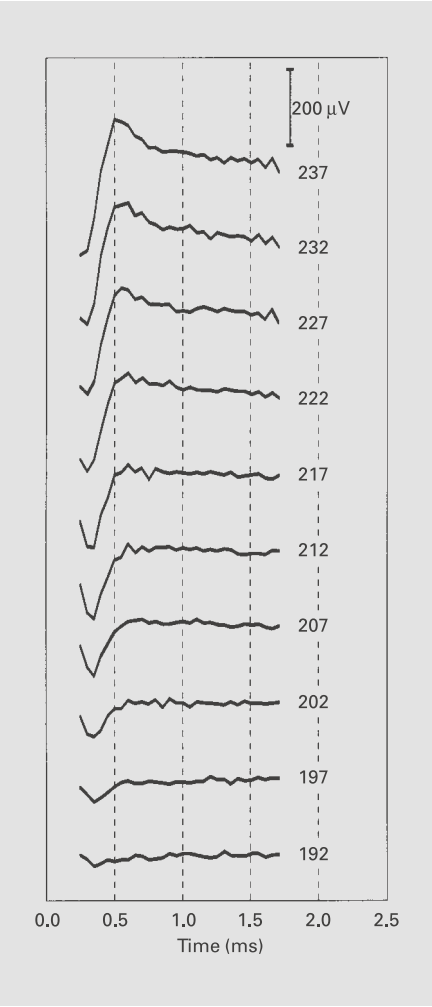

Fig. 6. NRT recording series illustrating how the response waveform changes from category Ic (no distinct P_1 peak) at lower stimulation levels (≤ 217 current level units) to category Ia (≥ 222 current level units) as the P_1 peak emerges at the higher stimulation levels, and eventually to category Ib responses (≥ 232 current level units) as the N_1 peak shifts towards shorter latencies with increasing stimulation level. The cascade display of response waveforms has been adjusted such that the end points of the waveforms are equally spaced for clarity.

high levels of synchronicity in the neural firing and/or to very fast conduction of the action potentials. It is conceivable that a very-short-latency N_1 peak occurring to the first phase of the biphasic stimulus pulse could appear during the presentation of the second phase of the stimulus pulse. It would then be impossible for any amplifier to actually capture such an N_1 peak embedded within the stimulus artifact itself.

Figure 6 shows an example from the NRT measurement data where the latency of the N_1 peak is observed to decrease with increasing stimulation level, to the extent that a visible N_1 peak at lower levels of stimulation disappears from the recordings at higher levels of stimulation. The decrease in the N_1 peak latencies could possibly be due to increased levels of synchrony in the firings of the auditory neurons at higher stimulation levels, leading to shorter integration times, and hence shorter latencies, of the total neural response.

The recording series shown in figure 6 also illustrates how category Ic responses could be seen to be related to category Ia responses. The P_1 peak is absent at lower levels of stimulation but gradually appears at higher levels. It is, however, not a general observation that category Ic responses changed to category Ia responses with increasing stimulation level. Instances of persistent category Ia or Ic responses with changing stimulation levels have also been recorded. Category Ic responses can also have an absent N_1 peak for the same reasons given above.

Double Peaks and Sites of Origin of the Neural Responses

Double peak ECAP waveforms recorded from an electrode placed on the auditory nerve as it exited the internal auditory meatus have been reported by Stypulkoswki and van den Honert [1984] from the cat. Double peak waveforms have also consistently been observed in the ECAP [Killian, 1994] (recorded using two electrodes placed in the auditory bulla) and also in wave I of electrically evoked auditory brainstem response recordings [Miller et al., 1993] (using far-field scalp electrodes) of guinea pigs. NRT data presented here show that such double peak waveforms are also to be found in NRT recordings using scala tympani electrodes from the human auditory periphery.

The results shown in figure 7 demonstrate that the two peaks do not necessarily grow at the same rate, implying that the two peaks represent two different and possibly independent components of the overall waveform. Note that the absence of a distinct N_1 peak at stimulation levels just above threshold compromises the accuracy of the response amplitude measurements. However, for these discussions, it is not so much the absolute amplitudes as the relative rate of growth of the two components that is of interest. When double peaks occur, the peak with the shorter latency is often only visible towards higher stimulation levels, as can also be seen from figure 7. One inter-

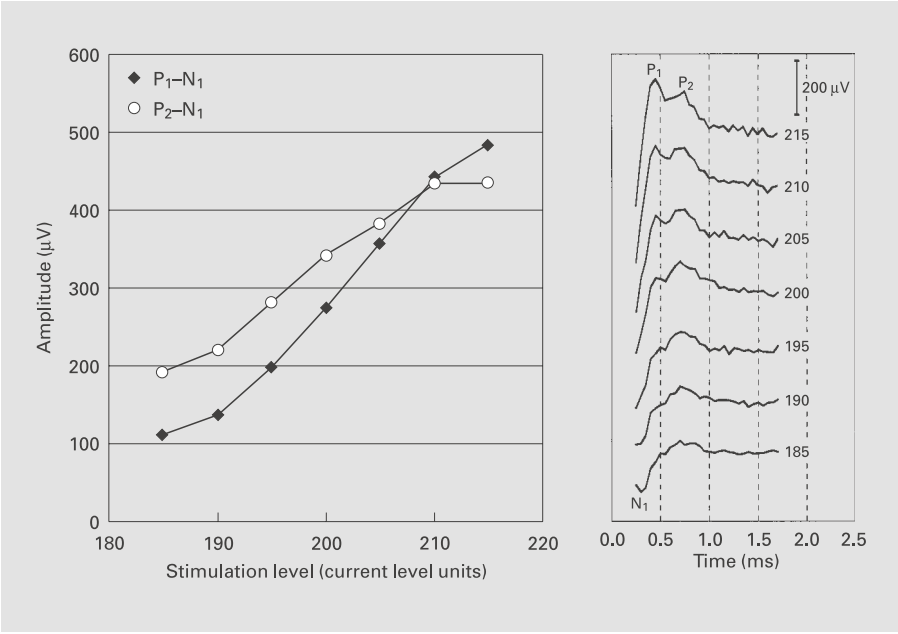

Fig. 7. NRT amplitude growth functions of the P_1 and P_2 peaks. Note also how the response waveform changes from category Ia to category II with increasing stimulation level.

pretation of this is that the two peaks represent responses from two different neural populations within the cochlea, one closer to the stimulation site and the other one more removed, and that the shorter latency peak which appears at higher stimulation levels corresponds to the more remote neural population which is stimulated by the enlarged electric field at higher stimulation levels. However, such an interpretation which assumed that only exactly two discrete populations were able to respond to the stimulus is simplistic. Extending this interpretation further to a more general case involving more than 2 discrete populations, or a more distributed population of excited neurons, would result in response waveforms with more than two peaks, or a broad peak that is smeared across a range of latency values. This is not the case as the double peak structure seen in figure 7 remains very distinct with increasing stimulation levels, and the likelihood of a compound response waveform arising from multiple sites of stimulation within the cochlea is diminished.

Another possible cause of the double peaks could be that they arise in reaction to the individual phases of the biphasic stimulus pulse. Such a possibility had been ruled out in earlier comparisons of recordings made using 25-μs and 50-μs phase width stimuli which both yielded responses whose two peaks were 300 μs apart, suggesting that the two peaks were independent of the phase duration of the stimuli. Furthermore, if the two phases of the stimuli were the cause of the double peaks, one would have expected to see these response waveforms more often than in the mere 9.5% of the cases reported here.

A more likely candidate for explaining the origin of these two components is that proposed by Stypulkowski and van den Honert [1984] based on ECAP recordings from the cat. According to their model, the earlier of the two peaks is attributed to direct excitation of the axons proximal to the soma and the later peak to activation of the dendrites. The cat data showed ECAP waveforms with two positive peaks at around 350-μs and 550-μs latencies. The waveforms and the order of magnitude of the reported latencies correlate well with the observed NRT waveforms. Further investigations [van den Honert and Stypulkowski, 1984] using single fiber recordings revealed a more complex pattern of activity with poststimulus time histograms, implying that the simple mechanism with two discrete sites of neural response activity may need to take into consideration other possible factors such as a continuous range of propagation delays as well as distributed excitation sites along the stimulated neurons.

Stypulkowski and van den Honert [1984] also reported that changes to the response waveform could be observed when small changes were made in the position of the intracranial recording electrode. This raised the possibility that the double peaks were merely an artifact of the location of the recording electrode. If this were so, one

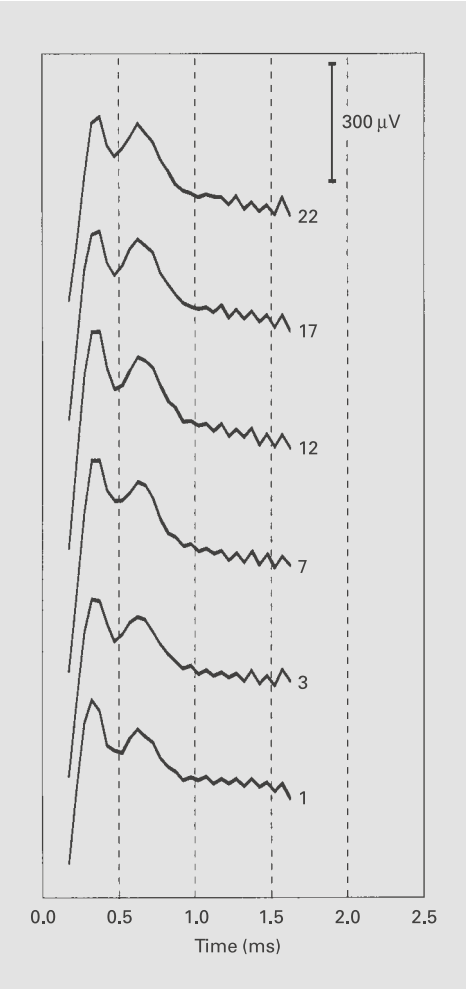

Fig. 8. Recording series with monopolar (MP1) stimuli were presented at electrode 10 (25 µs phase width, 224 current level units) and recorded at 6 different sites (electrodes 1, 3, 7, 12, 17, 22) along the array in the monopolar (MP2) configuration. The NRT waveforms show slight differences in response amplitude but no changes in the response waveform category as a function of recording site.

would expect similar changes in the NRT waveforms when different recording sites were used. This is not the case, as shown in figure 8 where the double peak structure was seen to be robust with respect to changes in the recording site both apical as well as basal of the stimulation site. Although one cannot be conclusive when comparing the effects of intracranial versus intracochlear recording sites in this manner, the NRT data support the view that the double peaks are more likely a result of the dendritic/axonal model than simply a byproduct of changing the recording site.

Implications of the Simulation Results

The simulation results show that the apparently trivial act of combining two single peak components to simulate the double peaks does not only give rise to double peak waveforms. Instead, depending on the relative amplitudes and onset times of the two components, the different types of NRT waveforms exhibited in our data could be reproduced. The success in modeling the different neural response waveform categories encourages the view that the neural response is produced by two components. The mathematical model suggests that these two components could be related to active axonal and dendritic processes, respectively.

The simulations of category Ia (and Ib) responses did not indicate conclusively whether only one or two components were present in the response waveform. Given that cochlear implant patients generally have a history of sensorineural hearing loss, it is expected that the majority of them would exhibit significant retrograde degeneration of their auditory peripheral processes in the manner indicated by physiological data from human temporal bone studies (e.g. Pollak and Felix [1985], Nadol [1990]). Other physiological studies (e.g. Goycoolea et al. [1990]) have suggested that the spiral ganglion cells may survive although their corresponding dendrites have degenerated. The fact that category Ia (and Ib) responses also represent the majority of observed neural response waveforms therefore suggests that these response waveforms are more likely to consist of only a single $r_a(t)$ axonal component. This, however, implies that the expected latency of the single positive peak would be shorter than that observed for category Ia (and Ib) responses. Miller et al. [1993] have suggested that the absence of the shorter latency positive peak in the resultant neural response could involve the excitation of predominantly high-capacitive membranes. The extra capacitance would then account for some of the increased response latencies.

The simulation results, however, also raised the possibility of two components adding up to yield a compound response waveform with a single positive peak. In this case, the proportion of $r_a(t)$ and $r_d(t)$ components that gave rise to category Ia (and Ib) responses would involve an axonal component which is about twice as large as the dendritic component. It is not possible to ascertain from the NRT data alone whether the size of the components would be related to the respective population densities of the axonal or dendritic processes, or if it would be related to the proximity of the stimulating electrode to these processes. That a dendritic component is implied at all sug-

gests that there are some surviving dendritic processes. Although in general the peripheral processes are expected to be missing in cases of sensorineural hearing loss, as already argued above, there is also a great deal of variability in nerve survival which is probably dependent on etiology. For instance, human temporal bone studies have been reported (e.g. Felix et al. [1997]) which support the presence of intact nerve fibers in the osseous spiral lamina in patients with prolonged sensorineural hearing loss.

The simulation of the category Ic responses is less ambiguous and clearly indicates the presence of two components to account for the missing positive P_1 peak. The model results suggest that these response waveforms consist predominantly of the axonal $r_a(t)$ component, but with a small dendritic component sufficient to cause the cancellation of the positive P_1 peak in the compound response waveform. The simulation results do not indicate whether the small dendritic component is a result of a small population of surviving dendritic processes or whether it is due to a more perimodiolar position of the stimulation electrode, resulting in stronger stimulation of the axonal processes compared to the dendritic processes.

Modeling of the category II responses with the double positive peak suggests a longer delay between the axonal and the dendritic components compared to the category I responses. This can be interpreted in two ways. On the one hand, the longer delay may be attributed to the dendritic component, implying that the dendritic activity has taken longer to propagate. The delay in the conduction of the dendritic component is attributed to longer peripheral processes as the neural conduction velocity is the more likely mechanism regulating propagation delays. The density of surviving dendrites is more likely to be reflected in the amplitude of the action potentials, but this is more difficult to quantify. The presence of longer and therefore more intact peripheral processes indicates a better likelihood of good dendritic survival. This is consistent with the observation that category II responses were significantly more often to be found towards the apical end of the array, where one would expect better survival of the dendrites. On the other hand, the increased delay between the two components can also be interpreted as an axonal component which occurs earlier. The net result is again a larger relative delay between the two components. As the scala tympani gets narrower towards the apex, the electrode array is constricted towards a more perimodiolar position. The resultant stimulation can then be expected to stimulate the axonal processes more intensely, resulting in axonal activity with shorter latencies. A dendritic component is still required to account for the later positive peak. Both interpretations indicate the presence of axonal and dendritic components, and either one or both mechanisms would give rise to the two components postulated by the model required to obtain a double peak response.

The present model involving only two components arising from within the same neural population is further supported by finite-element modeling and simulation results observed by F. Rattay, P. Lutter and H. Felix [unpubl. data], who were also able to demonstrate neural activity with two peaks having latencies similar to those measured using NRT.

The simulation results from such a simple model, however, still need to be interpreted with care. As it stands, the model provides only a first-order approximation of the neural mechanisms giving rise to the measured response waveforms. Factors that can influence the amplitudes and latencies of the potentials that contribute to the measured response waveform such as (1) different population densities (i.e. survival) of axonal and dendritic processes or (2) the proximity of the stimulating electrodes to these responding processes are not specifically built into the model. Furthermore, the simulation results are not able to conclusively identify whether there is only one or two components present, as in the case of category Ia and Ib responses. Van den Honert and Stypulkowski [1984] also suggested that action potentials could arise from a continuous progression of stimulation sites along the dendrite to the axon. Such a mechanism is not accounted for by the model in its present form.

Further studies to examine these correlations would be needed to refine the diagnostic power of such a simple mathematical model. It would be instructive to examine how the neural response waveforms differ when different electrode configurations are employed, particularly with the upcoming generation of modiolus-hugging electrode arrays which are expected to produce more localized stimulation sites. This will have an impact on the design of coding strategies and stimulation paradigms.

Clinical Implications

According to the simulation results, the shape (or waveform category) of the recorded neural response waveform provides some information regarding the amount of axonal versus dendritic activity. Double positive peak category II response waveforms appear to indicate the presence of two components, as do the 'no apparent positive peak' category Ic responses which require the

presence of two components to account for the 'missing' P_1 peak. Single positive peak category Ia and Ib response waveforms could be a result of either one or two components, although the likelihood is greater that these responses consist of only a single component. In terms of nerve survival, this would correspond to the more generally expected cases of retrograde degeneration. Additional information will be required to determine exactly whether these response waveforms consist of only one or two components.

The mathematical model therefore suggests that category Ic and category II responses indicate some degree of dendritic activity, while category Ia and Ib responses are more likely to correspond to poor dendritic activity. Further studies will have to be conducted to investigate the information-carrying capacities of these two modes of neural activity, and whether these indeed have any effect on sound perception. If an effect can be established, this will have very important implications for designing optimized coding strategies for individual patients. For instance, the response waveform category may be useful for the selection or exclusion of particular electrodes based on the expected information-carrying capacity of the active neural elements. Another possibility is that the response waveform may provide clues about the capacity of the active neurons in transmitting rate-coded information such that one kind of response waveform indicates high-rate coding strategies while another indicates lower-rate coding strategies. This kind of applications of objective NRT data in speech processor programming is particularly useful in situations where the patient, such as a child, is not able to give reliable feedback about loudness and quality judgements.

In addition to the clinical applications of NRT arising from the mathematical modeling results presented above, NRT is potentially a very useful clinical tool. The utility of the neural response measurement threshold as a guideline for fitting the maps of a patient's speech processor has been investigated by Franck [1999]. The use of the neural responses' amplitude growth function for electrode selection or exclusion in conjunction with an evaluation comparing different coding strategies is presently being examined in a number of clinical studies.

Conclusion

A simple mathematical model which linearly combines two-component waveforms was able to simulate the different NRT waveform categories obtained from 60 patients. The model was based on the assumption that the two components with different amplitudes and latencies originate either from the dendritic processes or from the axonal processes. Such a mechanism was postulated by Stypulkowski and van den Honert [1984] based on compound action potential recordings from the cat, and further indications of this mechanism are provided by the NRT measurements which demonstrate double positive peaks. Simulation results using the mathematical model described here support the hypothesis that such a mechanism is indeed in operation and that this mechanism not only gave rise to the double peak responses, but indeed also to the other types of NRT response waveforms observed. The simulation results also help provide important clues about the relative amount of axonal versus dendritic activity by breaking down a response into these two components, respectively. Double peak response waveforms appear to indicate good dendritic survival while single peak response waveforms suggest poor dendritic survival.

The implications of the relative amount of axonal versus dendritic activity for speech and sound perception remain to be investigated in future studies. This information would not only be useful in assisting clinicians in selecting an optimal stimulation paradigm for individual cochlear implant users, but could also have important ramifications for future electrode designs and the further development of coding strategies.

Acknowledgements

This work was supported by the Swiss National Foundation (Grant No. 3200-047325.96) and a research grant from Cochlear AG, Basel.

References

Abbas PJ, Brown CJ, Shallop JK, Firszt JB, Hughes ML, Hong SH, Staller SJ: Summary of results using the nucleus CI24M implant to record the electrically evoked compound action potential. Ear Hear 1999;20:45–59.

Boex C, Pelizzone M, Montandon P: Speech recognition with a CIS strategy for the Ineraid multichannel cochlear implant. Am J Otol 1996;17: 61–68.

Brown CJ, Abbas PJ, Gantz B: Electrically evoked whole-nerve action potentials: Data from human cochlear implant users. J Acoust Soc Am 1990;88:1385–1391.

Cafarelli Dees D, George C, Stevenson F, Sheridan C, Haacke NP: Comparison of two cochlear implant speech processors in better versus poorer performers. Ann Otol Rhinol Laryngol Suppl 1995;166:258–260.

Carter PM, Fisher AR, Nygard TM, Swanson BA, Shepherd RK, Tykocinski M, Brown M: Monitoring the electrically evoked compound action potential by means of a new telemetry system. Ann Otol Rhinol Laryngol Suppl 1995;166:48–51.

Charlet de Sauvage R, Cazals Y, Erre JP, Aran JM: Acoustically derived auditory nerve action potential evoked by electrical stimulation: An estimation of the waveform of single unit contribution. J Acoust Soc Am 1983;73:616–627.

Felix H, Gleeson MJ, Pollak A, Johnsson LG: The cochlear neurons in humans; in Iurato S, Veldman JE (eds): Progress in Human Auditory and Vestibular Histopathology. Amsterdam, Kugler, 1997, pp 73–79.

Franck KH: The Electrically Evoked Whole-Nerve Action Potential: Fitting Applications for Cochlear Implant Users; PhD thesis, University of Washington, 1999.

Goycoolea MV, Stypulkowski P, Muchow DC: Ultrastructural studies of the peripheral extensions (dendrites) of type I ganglion cells in the cat. Laryngoscope 1990;100:19–24.

van den Honert C, Stypulkowski PH: Physiological properties of the electrically stimulated auditory nerve. II. Single fiber recordings. Hear Res 1984;14:225–243.

Kiefer J, Muller J, Pfennigdorff T, Schon F, Helms J, von Ilberg C, Baumgartner W, Gstoettner W, Ehrenberger K, Arnold W, Stephan K, Thumfart W, Baur S: Speech understanding in quiet and in noise with the CIS speech coding strategy (MED-EL Combi-40) compared to the multipeak and spectral peak strategies (Nucleus). ORL J Otorhinolaryngol Relat Spec 1996;58: 127–135.

Killian MJP: Excitability of the Electrically Stimulated Auditory Nerve; PhD thesis, University of Utrecht, 1994.

Lai WK: An NRT Cookbook. Basel, Cochlear AG, 1999.

Miller CA, Abbas PJ, Robinson BK: Characterization of wave I of the electrically evoked auditory brainstem response in the guinea pig. Hear Res 1993;69:35–44.

Nadol JB Jr: Degeneration of cochlear neurons as seen in the spiral ganglion of man. Hear Res 1990;49:141–154.

Pollak AM, Felix H: Histopathological features of the spiral ganglion and cochlear nerve in temporal bones from three patients with profound hearing loss. Acta Otolaryngol Suppl (Stockh) 1985;423:59–66.

Skinner MW, Clark GM, Whitford LA, Seligman PM, Staller SJ, Shipp DB, Shallop JK, Everingham C, Menapace CM, Arndt PL, Antognelli T, Brimacombe JA, Pijl S, Daniels P, George CR, McDermott HJ, Beiter AL: Evaluation of a new spectral peak coding strategy for the Nucleus 22 channel cochlear implant system. Am J Otol Suppl 1994;15:15–27.

Skinner MW, Fourakis MS, Holden TA, Holden LK, Demorest ME: Identification of speech by cochlear implant recipients with the Multipeak (MPEAK) and Spectral Peak (SPEAK) speech coding strategies. I. Vowels. Ear Hear 1997;17: 182–197.

Stypulkowski PH, van den Honert C: Physiological properties of the electrically stimulated auditory nerve. I. Compound action potential recordings. Hear Res 1984;14:205–223.

Vandali A, Whitford LA, Plant K, Clarke GM: Speech perception as a function of electrical stimulation rate: Using the Nucleus 24 cochlear implant system. Ear Hear, in press.

Whitford LA, Seligman PM, Everingham C, Antognelli T, Skok MC, Hollow RD, Plant KL, Gerin ES, Staller SJ, McDermott HJ, Gibson WR, Clark GM: Evaluation of the Nucleus Spectra 22 processor and new speech processing strategy (SPEAK) in postlinguistically deafened adults. Acta Otolaryngol (Stockh) 1995;115: 629–637.

Wilson BS, Finley CC, Lawson DT, Wolford RD, Zerbi M: Design and evaluation of a continuous interleaved sampling (CIS) processing strategy for multichannel cochlear implants. J Rehabil Res Dev 1993;30:110–116.